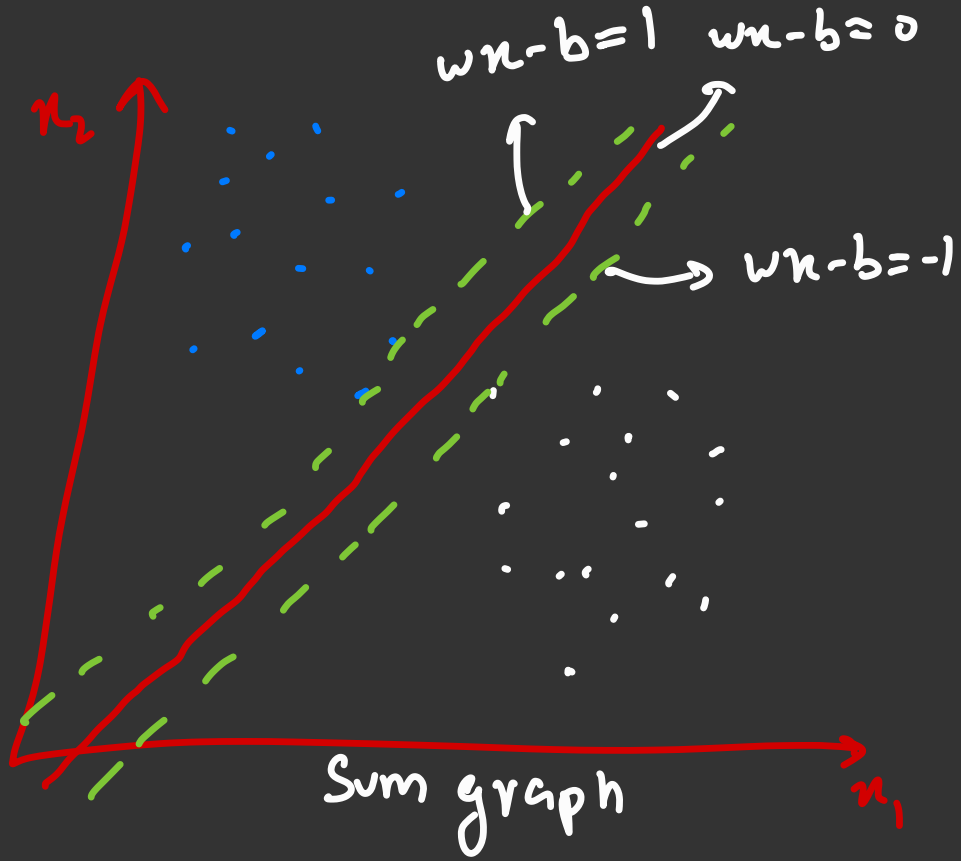
Support vector machine from scratch



* Red line is the hyperplane which defines the class

depending on where the point lies

Linear Model:

$wx - b = 0$

$wx - b \geq 1$ if $y_i = 1$

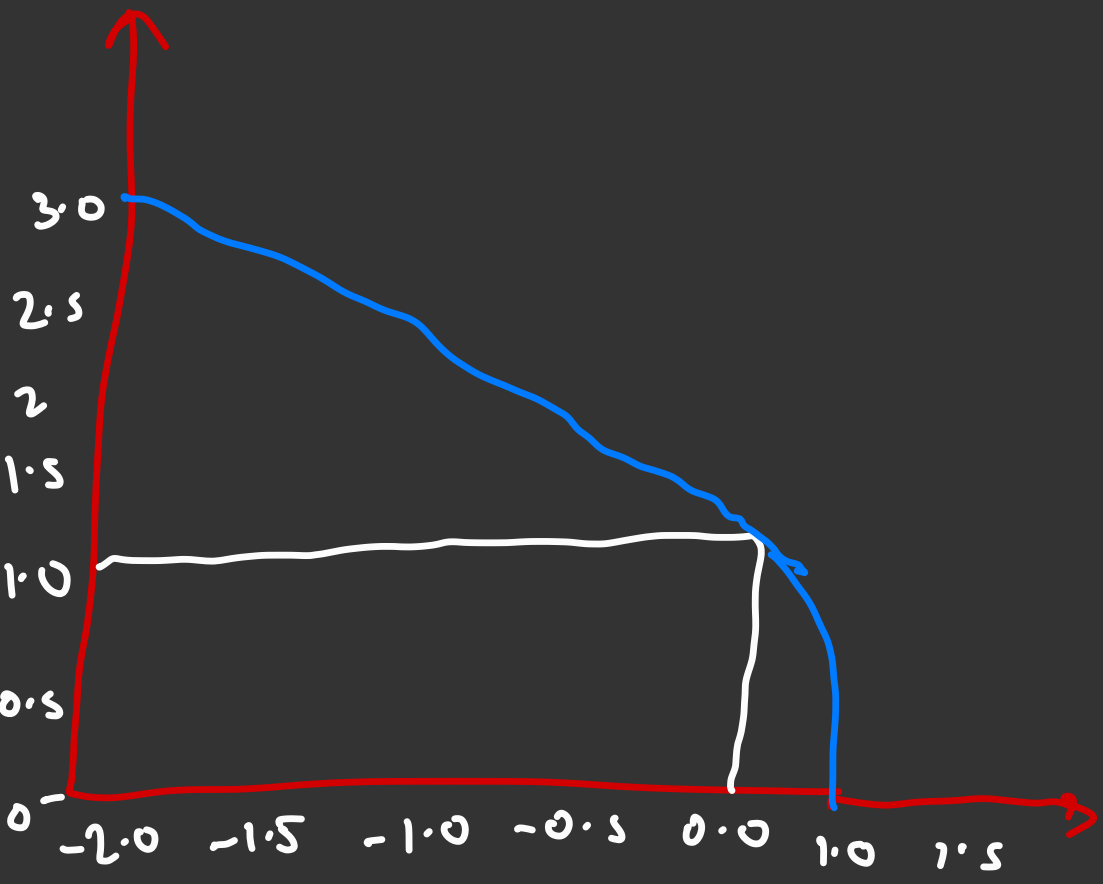
$wx - b \leq 1$ if $y_i = -1$

$y_i(wx - b) \geq 1$

Cost function

Hinge loss

$L = \max(0, 1 - y_i(wx - b))$



loss function graph

★ The further we are from the descision boundaries, higher

is the loss

Add Regularisation to minimise loss

$$J = \lambda \|w\|^2 + \underbrace{\frac{1}{n} \sum_{i=1}^{n} \max(0, 1 - y_i(wx_i - b))}_{\text{hinge loss}}$$

$\lambda \rightarrow$ parameter

$\|w\| \rightarrow$ magnitude of weight

if $y_i \cdot f(x) \geq 1$

$J_i = \lambda \|w\|^2$

else

$J_i = \lambda \|w\|^2 + 1 - y_i (wx - b)$

Gradients

if $y_i \cdot f(x) \geq 1$

$$\frac{dJ_i}{dw_k} = 2\lambda w_k$$

$$\frac{d J_i}{db} = 0$$

else

$$\frac{d J_i}{d w_k} = 2\lambda w_k - y_i \cdot x_i$$

$$\frac{d J_i}{db} = y_i$$

Update rule:

$$w = w - \alpha \, dw$$

$$b = b - \alpha \cdot db$$